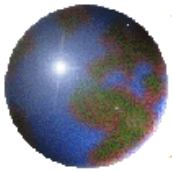


Global Positioning System (GPS)

Joe Montana

George Mason University

Source Material:

- http://www.trimble.com/gps
- Leila Z. Ribeiro Class Handouts

GPS Creation

- The U.S. Department of Defense decided that the military had to have a very precise form of worldwide positioning (began 1973 and fully operational 1995).
- And fortunately they had the kind of money ($12 Billion!) it took to build it.

What is GPS

- Worldwide radio-navigation system formed from a constellation of 24 satellites and their ground stations.
- Uses satellites as reference points to calculate positions accurate to a matter of meters (advanced forms of GPS can achieve centimeter accuracy).
- GPS receivers miniaturized and becoming very economical and accessible to the end users.
- Applications in cars, boats, planes, construction equipment, mobile smartphones, etc.

GPS Satellites

- **Name:** NAVSTAR **Manufacturer:** Rockwell International
- **Altitude:** 10,900 nautical miles
- **Weight:**1900 lbs (in orbit)
- **Size:**17 ft with solar panels extended
- **Orbital Period:** 12 hours
- **Orbital Plane**: 55 degrees to equatorial plane
- **Planned Lifespan:** 7.5 years
- **Current constellation:** As of March 9, 2016, there were 31 operational satellites in the GPS constellation (12,IIR-7,IIR(M)-12,IIF)

Ground Control Stations

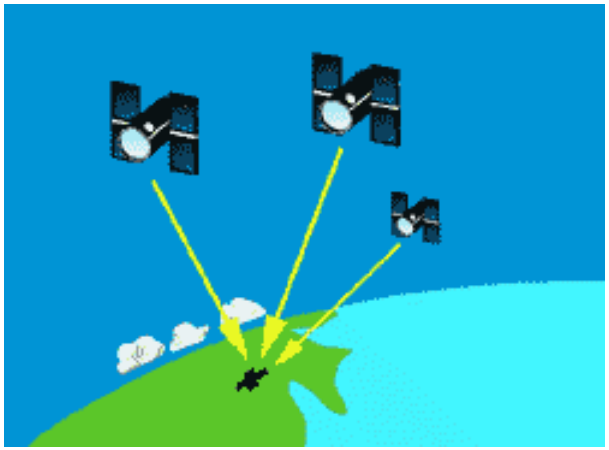
- Also known as the "Control Segment".
- Monitor the GPS satellites, checking both their operational health and their exact position in space.
- The master ground station transmits corrections for the satellite's ephemeris constants and clock offsets back to the satellites themselves.
- The satellites can then incorporate these updates in the signals they send to GPS receivers.
- There are five monitor stations: Hawaii, Ascension Island, Diego Garcia, Kwajalein, and Colorado Springs.

How GPS works

1. The basis of GPS is "**triangulation**" from satellites (formally speaking, “trilateration”).
2. To "triangulate," a GPS receiver **measures distance** using the travel time of radio signals.
3. To **measure travel time**, GPS needs very accurate timing which it achieves with some specific techniques.
4. Along with distance, the receiver needs to know exactly **where the satellites are in space**. High orbits and careful monitoring contribute to this accuracy.
5. Finally the receiver must **correct for any delays** the signal experiences as it travels through the atmosphere.

We will see each step next

1 - Triangulation from Satellites



- Use satellites in space as reference points for location on earth.
- How does the knowledge of distance from three (or more) satellites allow the position determination? …

Triangulation - Basics

1. Position is calculated from distance measurements (ranges) to satellites.

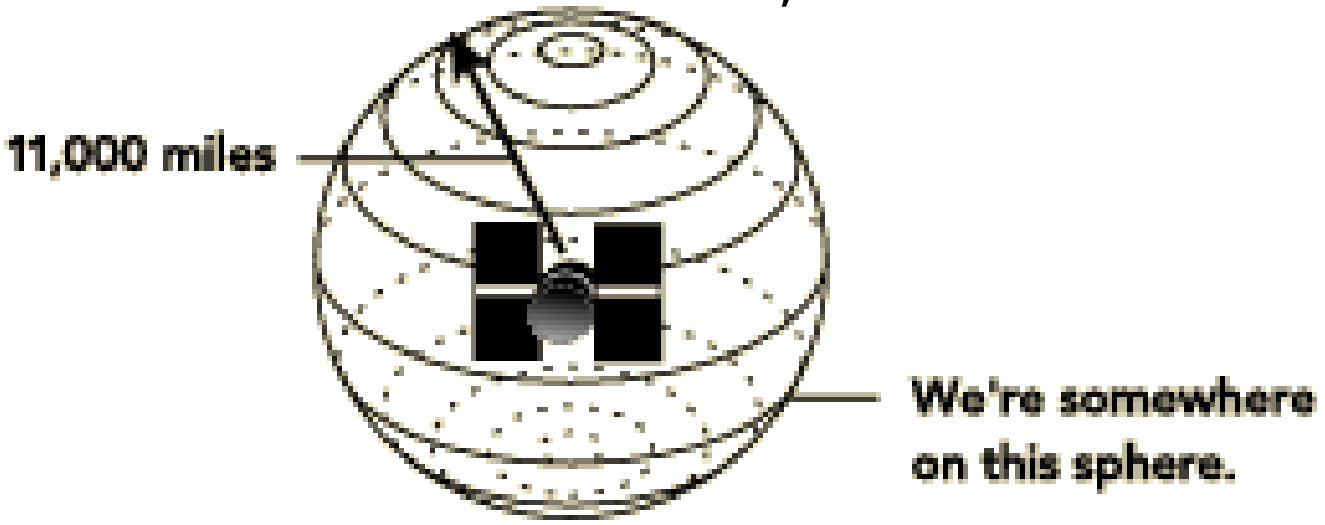
2. Mathematically we need four satellite ranges to determine exact position.

3. Three ranges are enough if we reject ridiculous answers or use other auxiliary.

4. Another range is required for technical reasons to be discussed later.

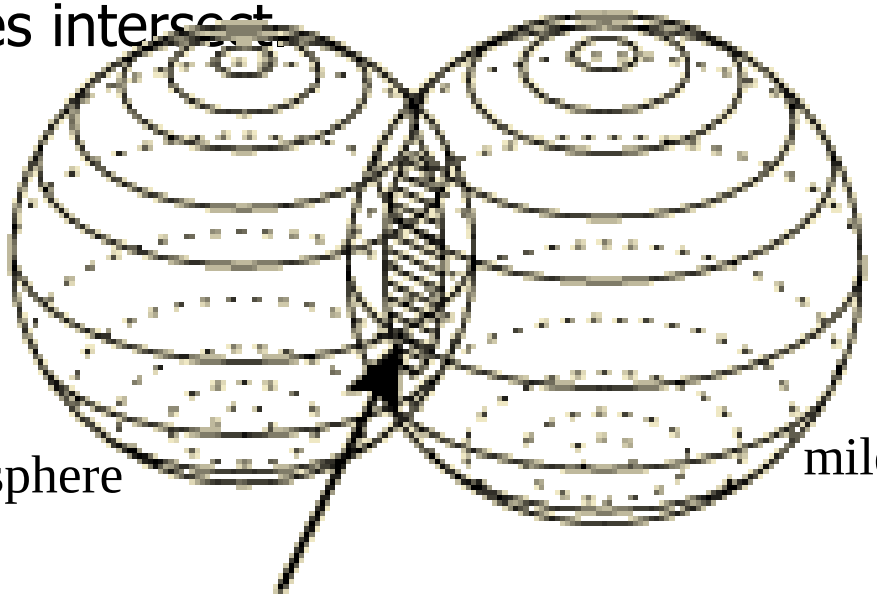
Distance to one satellite

- Suppose we measure our distance from a satellite and find it to be 11,000 miles. (How we measure that distance is the subject of further discussion)
- Knowing that we're 11,000 miles from a particular satellite narrows down all the possible locations we could be in the whole universe to the surface of a sphere that is centered on this satellite and has a radius of 11,000 miles.



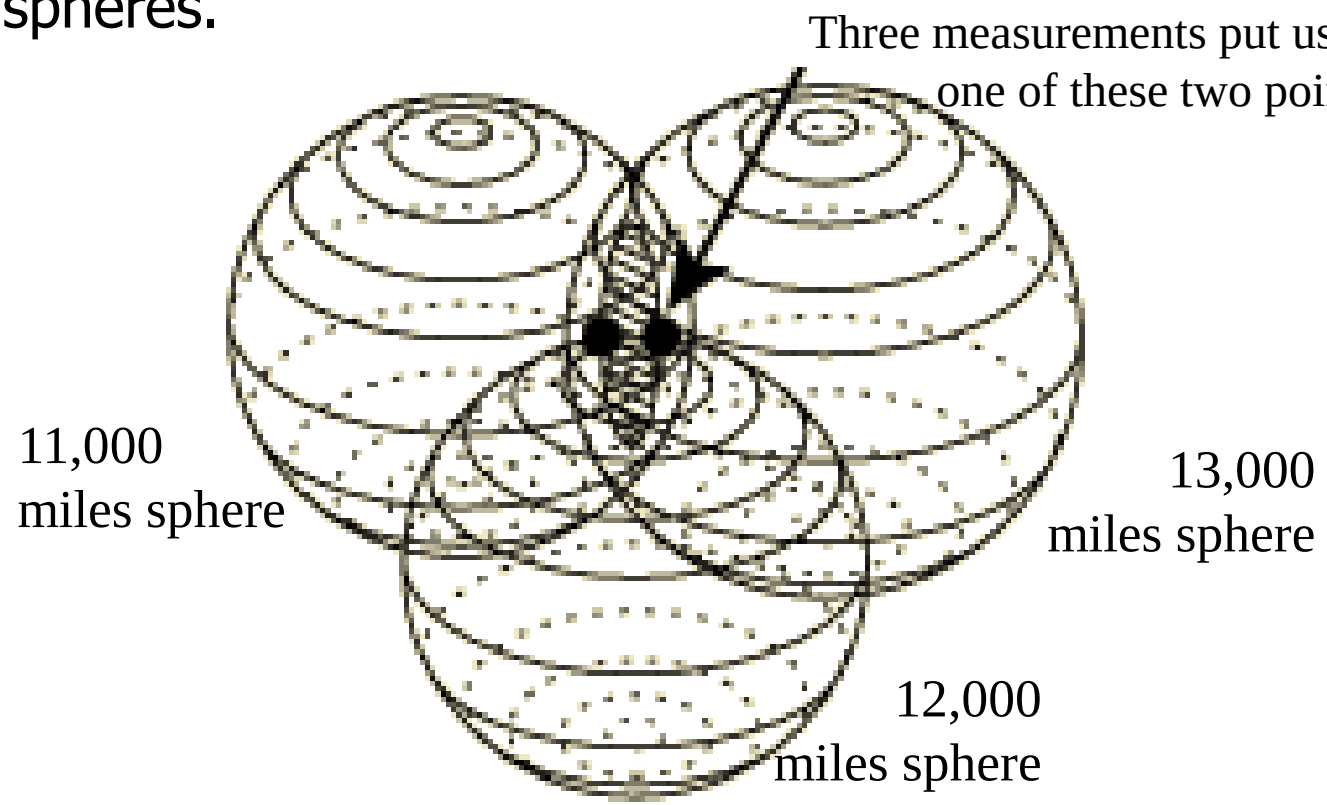
Distance to two satellites

- Next, suppose we measure our distance to a second satellite and find out that it's 12,000 miles away.
- That tells us that we're not only on the first sphere but we're also on a sphere that's 12,000 miles from the second satellite. Or in other words, we're somewhere on the circle where these two spheres intersect.



Distance to three satellites

- If we then make a measurement from a third satellite and find that we're 13,000 miles from that one, that narrows our position down even farther, to the two points where the 13,000 mile sphere cuts through the circle that's the intersection of the first two spheres.



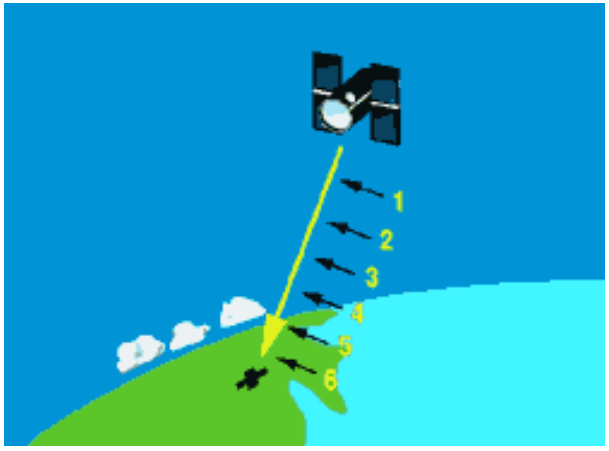
Triangulation - Summary

- By ranging from three satellites we can narrow our position to just two points in space.
- To decide which one is our true location we could make a fourth measurement. But usually one of the two points is a ridiculous answer (either too far from Earth or an impossible velocity) and can be rejected without a measurement.
- A fourth measurement does come in very handy for another reason however, but we will see that later.
- Next we'll see how the system measures distances to satellites.

2 - Measuring distance from a satellite

- From last section: position is calculated from distance measurements to at least three satellites. But how to measure the distance?

  Solution: By timing how long it takes for a signal sent from the satellite to arrive at the receiver.

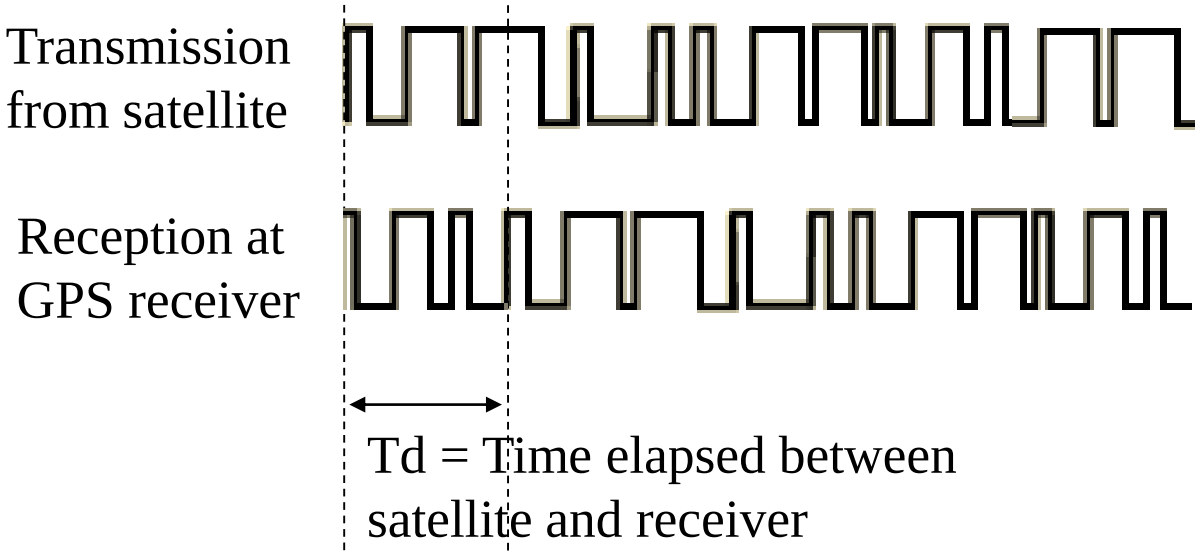


- Speed of light: c = 300,000 km/sec
- Distance to satellite is $d = c \times T_d$

The problem is measuring the travel time.

Measuring Travel Time

- A Pseudo Random Code (PRC) is transmitted from each satellite.
- Physically it's a pseudo-random sequence of "on" and "off" pulses.
- Receiver knows the time of transmission of the satellite sequence.
- By synchronizing the received sequence with a locally generated sequence, the receiver can identify the relative delay between the satellite and its location.



Reasons for using pseudo random sequences

- Avoid accidental synchronism with other interfering signal. The patterns are so complex that it's highly unlikely that a stray signal will have exactly the same shape.
- Since each satellite has its own unique Pseudo-Random Code they allow satellite identification. So all the satellites can use the same frequency.
- Pseudo-random sequences also make it more difficult for a hostile force to jam the system. In fact the Pseudo Random Code gives the DoD a way to control access to the system.
- Most importantly, the spread-spectrum effect gives spreading gain, which allows the receiver to “amplify” the signal at de-spreading. This enhances the link budget and allows economical GPS receiver (portable units with low gain antennas).

GPS Signals

- The GPS satellites transmit signals on two carrier frequencies.
  - The L1 carrier is 1575.42 MHz and carries both the status message and a pseudo-random code for timing.
  - The L2 carrier is 1227.60 MHz and is used for the more precise military pseudo-random code.

- Navigation Message: low frequency signal added to the L1 codes that gives information about the satellite's orbits, their clock corrections and other system status.

Pseudo-Random Codes

There are two types of pseudo-random code.

- The first pseudo-random code is called the C/A (Coarse Acquisition) code. It modulates the L1 carrier. It repeats every 1023 bits and modulates at a 1MHz rate. Each satellite has a unique pseudo-random code. The C/A code is the basis for civilian GPS use. CA code is at 1.024 Mbps.
- The second pseudo-random code is called the P (Precise) code. It repeats on a seven day cycle and modulates both the L1 and L2 carriers at a 10MHz rate. This code is intended for military users and can be encrypted. When it's encrypted it's called "Y" code. Since P code is more complicated than C/A it's more difficult for receivers to acquire. That's why many military receivers start by acquiring the C/A code first and then move on to P code. P code is at 10.24 Mbps.

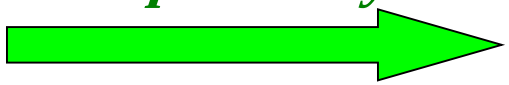
Summary – Measuring Distances

1. Distance to a satellite is determined by measuring how long a radio signal takes to reach the user from that satellite.
2. To make the measurement we assume that both the satellite and the user's receiver are generating the same pseudo-random codes at exactly the same time.
3. By comparing how late the satellite's pseudo-random code appears compared to the receiver's code, the receiver determines how long the signal took to reach it.
4. Multiply that travel time by the speed of light and you've got distance.

Summary – Measuring Distances

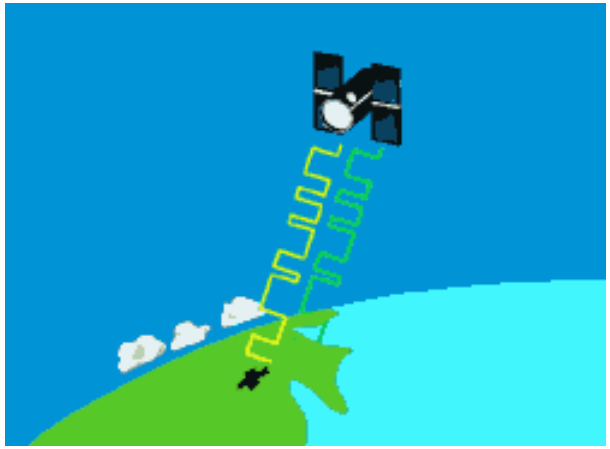
1. Distance to a satellite is determined by measuring how long a radio signal takes to reach the user from that satellite.
2. To make the measurement we assume that both the satellite and the user's receiver are generating the same pseudo-random codes at exactly the same time.
3. By comparing how late the satellite's pseudo-random code appears compared to the receiver's code, the receiver determines how long the signal took to reach it.
4. Multiply that travel time by the speed of light and you've got distance.

But to measure the time a perfect synchronism would be required!!

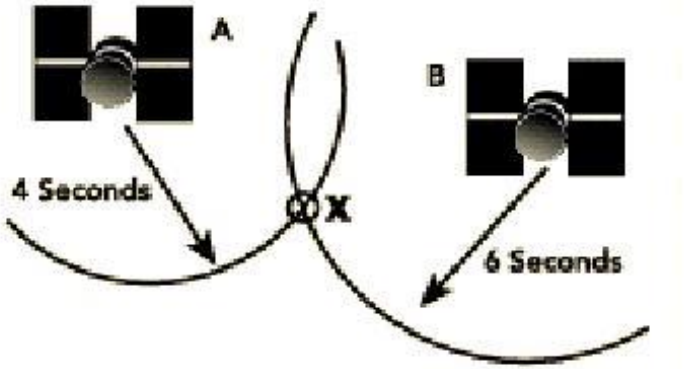


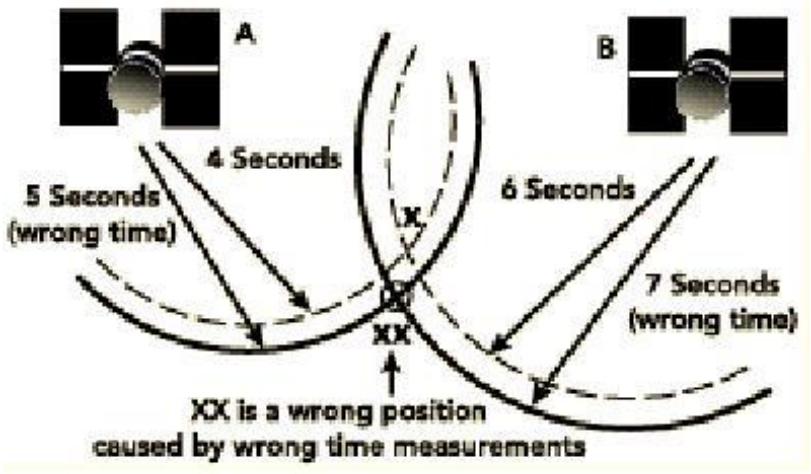
3 - Timing

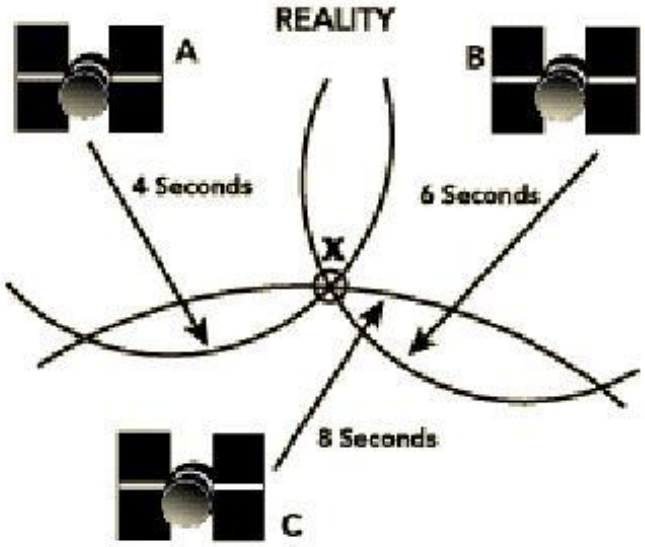
- Timing is critical: 1ms means a 200 mile error!
- Remember that both the satellite and the receiver need to be able to precisely synchronize their pseudo-random codes to make the system work.
- On the satellite side, timing is almost perfect because they have incredibly precise atomic clocks on board.
- But what about receivers on the ground?

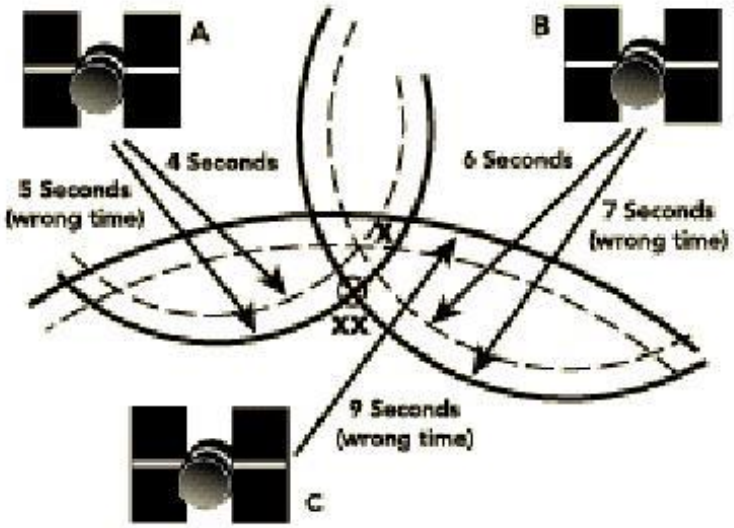


Position error due to wrong timing









Timing at receivers

- If our receivers needed atomic clocks (which cost upwards of $50K to $100K) GPS would be non-economical.
- Solution to this problem is to make an *extra* satellite measurement.
- This is one of the key elements of GPS and as an added side benefit it means that every GPS receiver is essentially an atomic-accuracy clock.
- In other words: if three perfect measurements can locate a point in 3-dimensional space, then four *imperfect* measurements can do the same thing.

How timing works at receivers

- If timing was perfect (i.e. if receiver's clocks were perfect) then all satellite ranges would intersect at a single point (which is the receiver's position). But with imperfect clocks, a fourth measurement, done as a cross-check, will NOT intersect with the first three.
- So the receiver's computer can detect the discrepancy in time measurements and recognize that it is out of synchronism with universal time.
- Since any offset from universal time will affect all of receiver measurements, the receiver looks for a single correction factor that it can subtract from all its timing measurements that would cause them all to intersect at a single point.
- That correction brings the receiver's clock back into sync with universal time, providing atomic accuracy time to it.

How timing works at receivers (cont.)

- Once receiver has the timing correction it applies to all the rest of its measurements and allows precise positioning.
- One consequence of this principle is that any GPS receiver will need to have at least four channels so that it can make the four measurements simultaneously.

•But for the triangulation to work we not only need to know distance, we also need to know exactly where the satellites are.

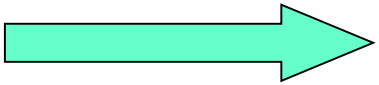
•In the next section we'll see how we accomplish that.

Summary - Timing

1. Accurate timing is the key to measuring distance to satellites.
2. Satellites are accurate because they have atomic clocks on board.
3. Receiver clocks don't have to be too accurate because an extra satellite range measurement can remove errors.

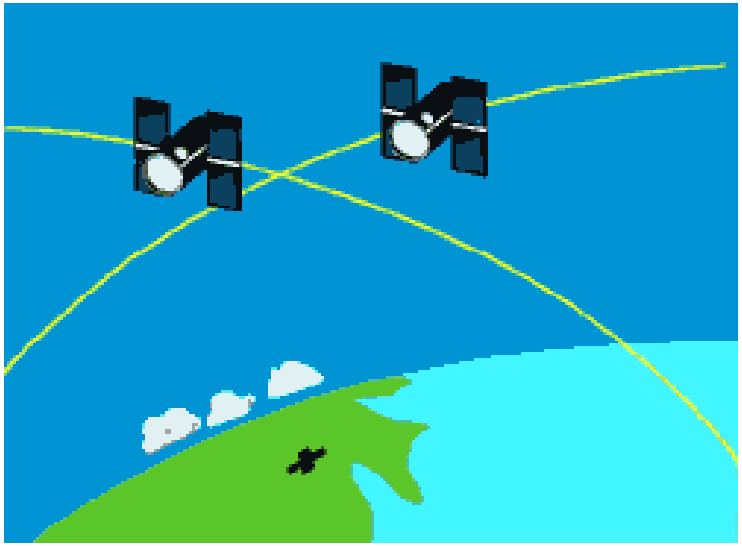
But for the triangulation to work we need not only to know distance, we also need to know exactly where the satellites are.

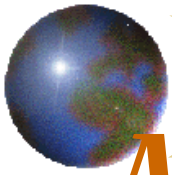
NEXT SECTION



4 - Satellite Position in Space

- On the ground all GPS receivers have an “almanac” programmed into their computers that tells them where in the sky each satellite is, moment by moment.





Monitoring Satellite Position

- Orbits constantly monitored by the Department of Defense.
- They use very precise radar to check each satellite's exact altitude, position and speed.
- Errors in position caused by gravitational pulls from the moon and sun and by the pressure of solar radiation on the satellites.
- The errors are usually very slight because of high orbit (MEO), but for accuracy they must be taken into account.

Monitoring Satellite Position (cont.)

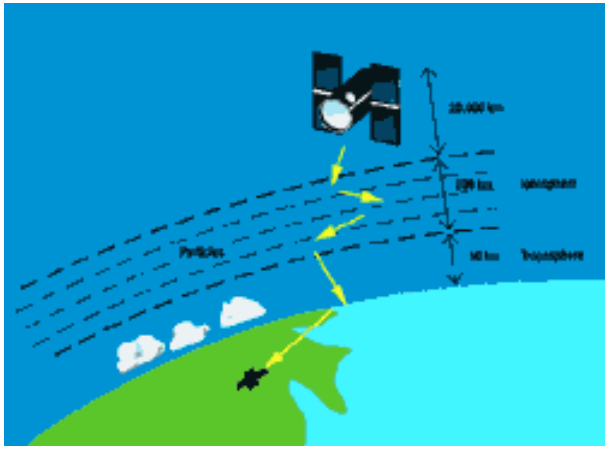
- Once the DoD has measured a satellite's exact position, they relay that information back up to the satellite itself. The satellite then includes this new corrected position information in the timing signals it's broadcasting.
- That is why a GPS signal is more than just pseudo-random code for timing purposes. It also contains a navigation message with ephemeris information as well.

Summary – Satellite Position

1. To use the satellites as references for range measurements we need to know exactly where they are.
2. GPS satellites are being at high orbits (MEO), are very predictable.
3. Minor variations in their orbits are measured by the Department of Defense.
4. The error information is sent to the satellites, to be transmitted along with the timing signals.

5 – Additional Errors

- Assumption distance to a satellite can be calculated by multiplying a signal's travel time by the speed of light was simplified so far: **speed of light is only constant in a vacuum**.
- As a GPS signal passes through the charged particles of the ionosphere and then through the water vapor in the troposphere it gets slowed down, and this creates the same kind of error as bad clocks.



Correcting delay errors

- To minimize the errors described, one can predict what a typical delay might be on a typical day. This is called modeling and provides considerable improvement but with limitations because atmospheric conditions are rarely typical.

- Another technique to minimize on these atmosphere-induced errors is to compare the relative speeds of two different signals. This "dual frequency" measurement is very sophisticated and is only possible with advanced receivers:
  - Physics says that as light moves through a given medium, low-frequency signals get "refracted" or slowed more than high-frequency signals. By comparing the delays of the two different carrier frequencies of the GPS signal, L1 and L2, we can deduce what the medium (i.e. atmosphere) is, and we can correct for it.
  - Unfortunately this requires a very sophisticated receiver since only the military has access to the signals on the L2 carrier.

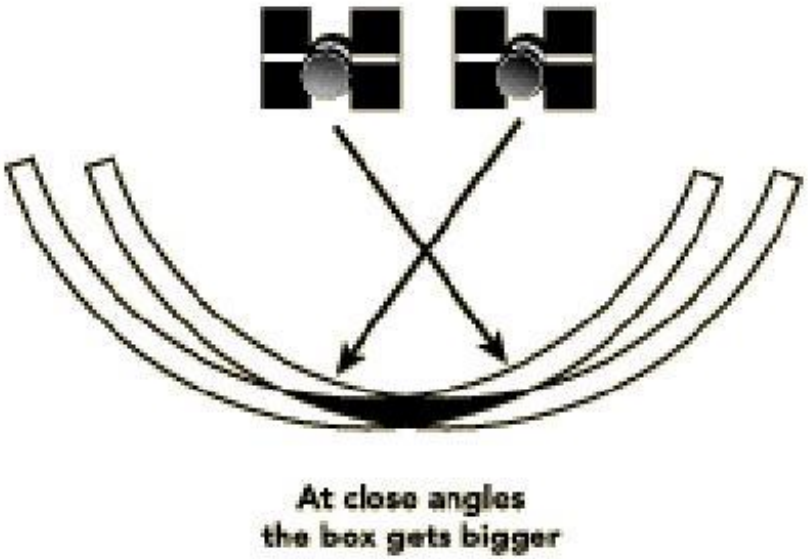
Other sources of error

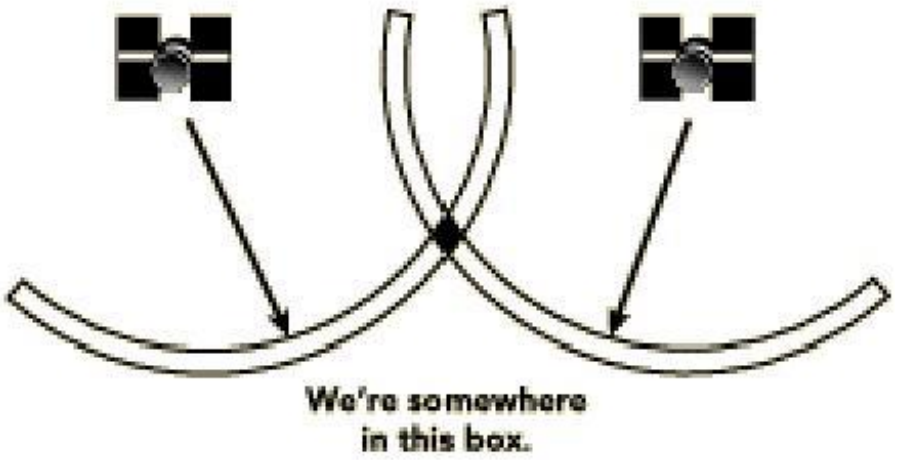
- Multipath error: The signal may bounce off various local obstructions before it gets to our receiver.
- Atomic clocks imperfections (small not null).
- Position detection errors.
- Intentional errors (removed in 2000) by the DoD. The policy was called "Selective Availability" or "SA" and the idea behind it was to introduce inaccuracies to make sure that no hostile force or terrorist group could use GPS to make accurate weapons.

Geometric Dilution of Precision

- Basic geometry itself can magnify these other errors with a principle called "Geometric Dilution of Precision" or GDOP.
- It sounds complicated but the principle is quite simple.
- There are usually more satellites available than a receiver needs to fix a position, so the receiver picks a few and ignores the rest.
- If it picks satellites that are close together in the sky the intersecting circles that define a position will cross at very shallow angles. That increases the gray area or error margin around a position.
- If it picks satellites that are widely separated the circles intersect at almost right angles and that minimizes the error region.
- Good receivers determine which satellites will give the lowest GDOP.

Geometric Dilution of Precision (cont.)





Summary - Correcting Errors

1. The earth's ionosphere and atmosphere cause delays in the GPS signal that translate into position errors.
2. Some errors can be factored out using mathematics and modeling.
3. The configuration of the satellites in the sky can magnify other errors.
4. Differential GPS can eliminate almost all error.

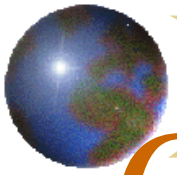
GPS Flavors

- **"Differential GPS,"** involves the use of two receivers. One monitors variations in the GPS signal and communicates those variations to the other receiver. The second receiver can then correct its calculations for better accuracy.
- **"Carrier-phase GPS"** takes advantage of the GPS signal's carrier signal to improve accuracy. The carrier frequency is much higher than the GPS signal which means it can be used for more precise timing measurements.
- **"Augmented GPS"** (aviation industry) involves the use of a geostationary satellite as a relay station for the transmission of differential corrections and GPS satellite status information. These corrections are necessary if GPS is to be used for instrument landings. The geostationary satellite would provide corrections across an entire continent.

GPS Accuracy

- C/A (civil): About 10 meters

- P (military): Can get down to centimeter with the use of differential GPS techniques.

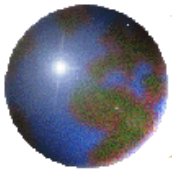


GPS Applications

- Civil Location - determining a basic position
- Tracking - monitoring the movement of people and things. Timing - providing atomic clock precision.
- Military: primary targeting and navigation system for US armed forces.
- Surveying: Mapping and locating land areas.
- Vehicular Navigation: on-car navigation systems.
- Ship navigation: Especially in coastal and inland waters.
- Aircraft navigations and landing: with development of Augmented GPS by FAA.

GPS Limitations

- Receiver must have line of sight to four or more satellites.
- Cannot work indoors of if sky is blocked (by buildings or other solid obstructions).
- Accuracy in vertical dimension is lower than in horizontal.
- CA code may be vulnerable to interference and jamming.



Other options of navigation systems

- **Landmarks:** Only work in local area. Subject to movement or destruction by environmental factors.
- **Dead Reckoning:**Very complicated. Accuracy depends on measurement tools which are usually relatively crude. Errors accumulate quickly.
- **Celestial:**Complicated. Only works at night in good weather. Limited precision.
- **OMEGA:**Based on relatively few radio direction beacons. Accuracy limited and subject to radio interference.
- **LORAN:**Limited coverage (mostly coastal). Accuracy variable, affected by geographic situation. Easy to jam or disturb.
- **SatNav:**Based on low-frequency doppler measurements so it's sensitive to small movements at receiver. Few satellites so updates are infrequent.